
Archival Featured Article

Adult Innumeracy: Are Students' Arithmetic Scores Affected by Gender-Related Stereotype Threat or IQ?

Viviane L. Falardeau¹, Lionel G. Standing¹

¹ Psychology Department, Bishop's University, Sherbrooke, Quebec, Canada

Correspondence: Lionel G. Standing (lstanding@ubishops.ca)

ABSTRACT – Based on past studies of the effect of stereotype threat upon test scores, this study examined the performance of female undergraduates ($N = 40$) on the D'Amore Test of elementary arithmetic, and the Wonderlic IQ test, after reading a stereotypically negative or positive gender-related prime. Overall, 67.5% of the subjects failed the D'Amore Test, a figure in agreement with the results of previous studies. No effect of the stereotyped priming articles was found on math scores, and no consistent correlation of D'Amore scores was found with IQ scores. No evidence was found that emotional or cognitive ability factors, rather than poor math skills, are responsible for failures on the D'Amore Test.

Keywords:
Stereotype threat;
Gender differences;
Academic performance;
Arithmetic skills;
Intelligence testing;
Cognitive ability;
Undergraduate students

Introduction

Some recent studies have required university students to take a test of elementary arithmetic, the D'Amore Test (DT), which uses extremely simple questions that were designed originally for third-grade students in 1932 (Cornwall, 1999; Hume, 1932). These items are shown in Appendix A. Surprisingly, it has been found that the majority of undergraduates fail to answer all of these ten questions correctly (Standing, 2006; Rustemeyer & Stoeger, 2007). These errors are widely distributed across the ten test questions (Standing et al., 2006).

Beyond replicating this finding, the present study was intended to look for possible emotional and/or intellectual ability factors that might contribute to the effect. The simplest explanation is in terms of subjects' defective basic math skills. However, it is increasingly recognized that

ISSN:
1541-745X (Print)
2169-3951 (Online)

Editor's note: This article is a reprint of an original work published in 2010 in *Psychology Journal*, Vol. 7(4), pp. 110-115.



©2010/2024 IDR Ltd Co. This is an open access article under the Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>).

emotional factors can affect mathematical reasoning (Else-Quest et al., 2008), and some aspects of intelligence in a broad sense have traditionally been seen as akin to the type of reasoning that is involved in mathematics.

Explicit prejudice is rare today, but judgmental attitudes based on race, age, or gender are still almost universal at an implicit level (Greenwald et al., 1998). Researchers have consequently been interested in the effect that stereotype priming may have on a person's performance, giving rise to the idea of a "stereotype threat." Although gender differences in performance on mathematical tasks are small and inconsistent (Else-Quest et al., 2010), reminding female subjects of their gender has been found to depress their performance on subsequent mathematical tasks (Ambady et al., 1999). Priming techniques are of interest here because they are based on the idea of activating a person's preconceived attitudes, without their conscious awareness. For these reasons, the present study examines the effect of stereotype priming on the mathematical performance of female students on the DT.

Considerable research has also been performed on the negative effect of stereotype threat upon IQ test performance. For example, African Americans do better on IQ tests if they are told they are participating in an experiment rather than providing a measure of their ability (Steele & Aronson, 1995). This indicates the possibility of gender stereotypes affecting IQ performance, and accordingly an IQ test was given to the participants here, subsequent to the DT, to investigate this effect. The Wonderlic Personnel Test of IQ was used, because it is conveniently short (12 minutes), with satisfactory validity and reliability (McKelvie, 1989).

It was hypothesized, in line with past research, that female subjects receiving a pro-female prime would do better on the DT than females given a pro-male prime. The participants were also expected to perform better on the Wonderlic IQ test if they had been positively primed, and a significant correlation was anticipated between the Wonderlic and DT scores. Finally it was predicted that scores on the DT would be low overall, in accord with previous findings.

Method

Participants

The forty female subjects of this study were European Canadian students at Bishop's University, from various programs. Their mean age was 21.84 years ($SD = 2.25$, range 18.00 to 29.83). Twenty subjects were randomly assigned to receive the pro-female priming article, and twenty the pro-male prime.

Materials

Two versions ("pro-female" and "pro-male") of an apparently realistic web article were prepared, to emphasize recent research findings pointing either to male or to female superiority in mathematical performance today. The articles were identical except that the words *male* and *female* were switched in each of five sentences which drew gender comparisons, e.g., "Hiklemann and Richards... found that there was a highly significant correlation between women's/men's poor performance on mathematical tasks and their level of dislike of math." (No italics in original). These two versions, each 313 words long, were adapted from a ScienceDaily article ("Women's Math Performance Affected by Theories on Sex Differences," 2006).

The standard DT was employed, incorporating the ten items as listed in Appendix A, as well as the Wonderlic Personnel Test (Wonderlic, 1983), plus consent forms and personal information forms.

Procedure

Participants were tested individually, in a quiet room. Having read and signed the consent form, they were handed a copy of the form to be filled in with basic personal information, one of the two versions of the article (randomly assigned), a blank sheet of paper for rough work, and a copy of the DT, all stapled together in that order. Subjects were instructed to fill out the personal information form and then carefully read the article. Only once they had read the article they completed the DT. No calculator was allowed, but they were told that they had as much time as they needed to finish the test. They were then wished good luck by the experimenter.

Following the DT, subjects were given the Wonderlic IQ test and asked to read the first page of instructions. Once they had done this, and confirmed that they had no questions, they were asked to begin and told that they would be stopped when the time was up.

After twelve minutes, the participants were asked to put their pen down and were thanked for their participation and debriefed on the goal of the study. To ensure that the subjects did not feel discomfort in completing the IQ test, they were informed that their scores would not be evaluated individually and used only for overall comparisons between conditions.

The design of this study was a between-subjects one-way ANOVA, the independent variable being the type of prime in the article (pro-female math performance priming article, or the pro-male prime). The dependent variables were the subject's scores on the DT and the Wonderlic test.

Results

The data were examined to see whether there was a difference between the mean scores on either test between the subjects in the two prime conditions, and whether these test scores were correlated. Possible demographic correlates were also examined. One subject's IQ score was eliminated as an outlier, since previous to her completion of the Wonderlic test this individual admitted having taken the same test a few days prior, resulting in an abnormally high Wonderlic score of 40 (3.87 *SD* above the mean).

Descriptive Statistics

Participants on average had completed four semesters at the university level; 42.5% were from the province of Quebec, while 40% were from another province of Canada and 15% were international students (one subject did not provide information on where she was from). The majority of students were enrolled in the departments of Psychology (45%), Fine Arts (20%), or the Pure and Applied Sciences (22.5%). The other 22.5% were in Business, Education, or another social science other than Psychology. Most of the subjects (75%) had been enrolled in University level math courses.

Overall, 67.5% of subjects failed the DT (75% and 60%, respectively, for the pro-female and pro-male prime groups), where failure is defined as a score below 10 items correct. For the pro-female and pro-male prime groups, respectively, the mean scores on the DT were 8.53 and 8.73 (*SD* = 1.33 and 1.39). For the Wonderlic test the corresponding means were 23.00 and 21.16 (*SD* = 5.13 and 3.93).

Effects of Prime on DT and Wonderlic Scores

Separate one-way ANOVAs indicated that neither the IQ scores nor the DT scores were affected by the type of priming article presented to the participant; $F(1, 38) = 1.57, p = .218$, and $F(1, 39) = 0.22, p = .645$, respectively.

Correlation Between Test Scores

Contrary to expectations, no overall correlation was found between subjects' DT and Wonderlic scores, $r(38) = .252, p = .121$. However, the group that received the pro-male prime did show a positive association between these two measures, $r(18) = .537, p = .018$, whereas the pro-female prime group did not, $r(18) = .069, p = .772$.

Demographic Correlations

Correlating the different academic divisions with test scores, it was found that students enrolled in the Education program had higher IQ scores, $r(38) = .352, p = .028$, whereas those in the Business department had lower math scores, $r(38) = -.348, p = .028$.

Concerning other demographic correlates, only subjects' age was linked to success on the DT; the older the subject, the better they did on the math test, $r(38) = .370, p = .019$. No other correlation was found with math scores, and no demographic correlate was found for the IQ test.

Discussion

The results confirm previous findings that a majority of undergraduates cannot reliably perform simple arithmetic and also indicate that the type of prime presented to the subjects did not affect either their math performance or their IQ scores. Although past research has consistently found a negative influence of gender-related stereotype threat, it can only have been of negligible importance here in comparison with the problem of math skill. However, some participants who were psychology students had possibly heard about stereotype threat, which may have decreased the primes' effect on their performance. Participants may not have read their article attentively enough, reducing its impact, or may have seen it as lacking in authenticity.

The overall mean scores on the DT found here are almost identical to those reported by Standing (2006). Even though three quarters of the present subjects had taken math courses at a reputable liberal arts college, 67.5% of them failed the test, since they scored below ten. This number is close to the values reported by Standing (2006), and by Standing et al. (2006), who found failure rates for the DT of 66.7% and 59.7%, respectively. However, even these figures are slightly better than the 74% failure rate for the DT reported by Rustemeyer and Stoeger (2007) with German students.

A moderate positive correlation between DT and Wonderlic scores was found, although only for the subjects receiving the pro-male prime. No obvious reason is evident for this discrepancy. However, this unreliable association suggests that failures on the DT are more likely to reflect deficiencies of mathematical training than of intrinsic cognitive ability.

Education students scored higher on the Wonderlic test. This may be because these students take English grammar and literature courses, giving them an advantage in those questions of the IQ test which involve word meaning. Moreover, in accordance with previous research (Standing, 2006), students in the Business program showed lower arithmetical performance than other students, even though they are required to have some mathematical knowledge. One possibility is that, because their studies require the constant use of a calculator, they have become so dependent on it that they are at a loss when its use is forbidden (Rustemeyer & Stoeger, 2007).

The results showed that the older the subject, the better their IQ score. This finding may be a cohort effect, as these older subjects passed through a different type of school system, or may be a true age effect. Increasing maturity and familiarity with testing may have better allowed them to take the pressure of being timed than the younger students, or age may foster the acquisition of all-round behavioral skills such as attentional focusing that could facilitate the completion of an IQ test.

We conclude from the present results that the most parsimonious explanation for the high level of failure on the DT remains that of subjects' poor basic math skills, primarily attributable to their prior instruction in the education system, rather than emotional factors or differences in intellectual ability.

Acknowledgements: The support of the BU Research Committee is gratefully acknowledged. Lou D'Amore is thanked for providing the arithmetic test bearing his name.

References

- Ambady, N., Pittinsky, T. L., & Shih, M. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science*, 10, 80–83. <https://doi.org/10.1111/1467-9280.00111>
- Cornwall, C. (1999, February). Why Johnny can't add. *Reader's Digest*, 154, 38–43.
- Else-Quest, N. M., Hyde, J. S., & Hejmadi, A. (2008). Mother and child emotions during mathematics homework. *Mathematical Thinking and Learning*, 10, 5–35. <https://doi.org/10.1080/10986060701820386>
- Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin*, 136, 103–127. <https://doi.org/10.1037/a0018053>
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74, 1464–1480. <https://doi.org/10.1037/0022-3514.74.6.1464>
- Hume, W. E. (1932). *The opportunity plan: Arithmetic, senior third class*. Toronto, Canada: Thomas Nelson & Sons.
- McKelvie, S. J. (1989). The Wonderlic Personnel Test: Reliability and validity in an academic setting. *Psychological Reports*, 65, 161–162. <https://doi.org/10.2466/pro.1989.65.1.161>
- Rustemeyer, R., & Stoeger, H. (2007). Does calculator use explain why university students cannot perform elementary arithmetic? *Psychological Reports*, 100, 1270–1272. <https://doi.org/10.2466/pro.100.4.1270-1272>
- Standing, L. G. (2006). Why Johnny still can't add: Predictors of university students' performance on an elementary arithmetic test. *Social Behavior and Personality*, 34, 151–160. <https://doi.org/10.2224/sbp.2006.34.2.151>
- Standing, L. G., Sproule, R. A., & Leung, A. (2006). Can business and economics students perform elementary arithmetic? *Psychological Reports*, 98, 549–555. <https://doi.org/10.2466/pro.98.3.549-555>
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, 69, 797–811. <https://doi.org/10.1037/0022-3514.69.5.797>
- Women's math performance affected by theories on sex differences. (2006). *ScienceDaily*. <http://www.sciencedaily.com/releases/2009/10/061019161245.htm>
- Wonderlic, E. P. (1983). *Wonderlic Personnel Test manual*. Libertyville, IL: Wonderlic, Inc.

Appendix A

The D'Amore Test of Elementary Arithmetic (from Hume, 1932)

Please complete the D'Amore Test below, in writing. [No calculators are allowed, but you can write rough work on this form or the back.]

1. Subtract these numbers: $9,864 - 5,947$
2. Multiply: 92×34
3. Add the following: $\$126.30 + \$265.12 + \$196.40$
4. An airplane travels 360 kilometers in three hours. How far does it go in one hour?
5. If a pie is cut into sixths, how many pieces would there be?
6. William bought six oranges at 5 cents each and had 15 cents left over. How much had he at first?
7. Jane had \$2.75. Mary had 95 cents more than Jane. How much did Jane and Mary have together?
8. A boy bought a bicycle for \$21.50. He sold it for \$23.75. Did he gain or lose and by how much?
9. Mary's mother bought a hat for \$2.85. What was her change from \$5?
10. There are 36 children in one room and 33 in another room in Tom's school. How much will it cost to buy a crayon at 7 cents each for each child?

Answers

1. 3,917
2. 3,128
3. \$587.82
4. 120 km
5. 6
6. 45¢
7. \$6.45
8. Gain \$2.25
9. \$2.15
10. \$4.83